

AD-A038 242

MICHIGAN UNIV ANN ARBOR SYSTEMS ENGINEERING LAB
APPROXIMATION BY BERNSTEIN SYSTEMS.(U)

F/G 12/1

UNCLASSIFIED

SEL-108

AFOSR-TR-77-0369

AF-AFOSR-3160-77
NL

1 OF 1
AD
A038242



END

DATE
FILMED

5-77

AEUSK - IR - 47 - 0369

3
NW

ADA 038242

THE UNIVERSITY OF MICHIGAN

SYSTEMS ENGINEERING LABORATORY ✓

Department of Electrical & Computer Engineering
College of Engineering

SEL Technical Report No. 108 ✓

APPROXIMATION BY BERNSTEIN SYSTEMS

March 1977

Project Director

William A. Porter

DDC
RECEIVED
APR 11 1977
A

DDC FILE COPY

Approved for public release;
distribution unlimited.

ACQUISITION FOR	
NTIS	WIDE SPREAD
DDC	WIDE SPREAD
UNANNOUNCED	
JUSTIFICATION	
BY	
EXTENSION/AVAILABILITY	
ALIAS	
A	

Prepared under

AFOSR Grant
AFSOF 770352 77-3160

VDV 03857S

COBA LIFE COBA

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DDC
This technical report has been reviewed and is
approved for public release IAW AFR 190-12 (7b).
Distribution is unlimited.
A. D. BLOSE
Technical Information Officer

APPROXIMATION BY BERNSTEIN SYSTEMS*

William A. Porter[†]

O. Abstract Let K denote an arbitrary compact subset of L_2 . Let f be any causal continuous function on L_2 . Then there is a linear differential system: $\dot{x}(t) = A(t)x(t) + B(t)u(t)$, and a memoryless polynomial state to output map $y(t) = \phi(x(t))$ such that the system, \hat{f} , thereby computed satisfies

$$\sup_{x \in K} ||f(u) - \hat{f}(u)|| < \epsilon$$

where $\epsilon > 0$ is arbitrary. This and other results are developed.

1. Introduction. This article deals with the approximation of nonlinear systems by polynomial operators. For perspective it is helpful to consider the familiar Volterra series expansion on L_2 given by

$$p(x) = k_0 + \int k_1(\alpha)x(\alpha)d\alpha + \iint k_2(\alpha,\beta)x(\alpha)x(\beta)d\alpha d\beta + \iiint k_3(\alpha,\beta,\gamma)x(\alpha)x(\beta)x(\gamma)d\alpha d\beta d\gamma + \dots \quad (1)$$

where the kernels k_0, k_1, \dots, k_n satisfy properties suitable to an operator on L_2 . For the obvious reasons we refer to each term on the right hand side as a power function. If the number of terms is finite then p is said to be a polynomial operator. Our interest in polynomial operators centers on their use as approximates of the more general nonlinear functions on L_2 .

* Supported in part by the United States Air Force Office of Scientific Research, Grant No. ~~77-0352~~ 77-2160

[†] Computer, Information, and Control Program, The University of Michigan, Ann Arbor, Michigan.

The relevant literature may be grouped into two subcategories. First there is the analytic theory in which f is assumed to have derivatives (Frechet or Gateau) of all orders and p arises as a power series expansion on a bounded domain. This line of development was initiated by Volterra [1] and was first applied in a systems setting by Weiner and in ensuing years several others including [2], [3], and [4]. More recently [5], [6], [7], [8] have investigated the Volterra expansion of solutions to nonlinear differential equations with current emphasis on computation and convergence problems.

The analytic theory identifies the power functions with the derivatives of the system function, f , to be approximated. The requisite differentiability is a severe condition, however, a fringe benefit accrued is that the causality of the power functions is dictated by the causality of f .

In an independent line of development the polynomic approximation problem has been approached as a generalization of the classic Weierstrass result. In this setting the function, f , to be approximated need not be differentiable. Emphasis is placed on uniformly approximating f , by the polynomial p , over an arbitrary compact set. In this approach the causality issue is less trivial. The computation of the power functions, which no longer represent derivatives, has here to fore been obscure. In sections (3) and (4) we develop the concept of a Bernstein differential system. This system provides one constructive realization of the Weierstrass approach. In section (5) we

compare in part the properties of the analytic theory, the Weierstrass based theory and in particular the Bernstein system realization of section (4).

An incidental bonus of the Bernstein system is a pseudo sampling theorem for systems. In short, given an arbitrary system with a prescribed continuity modulus, the density with which one must input-output sample in order to be able to approximately reconstruct the system is established.

For efficiency of presentation we shall in many cases be overly restrictive in the assumptions made, for example we consider only Hilbert spaces. Also in sections (3) and (4) the development is purposely constrained so as to use classical results on the Bernstein polynomials.

In closing this introduction it is noted that the admittance characteristic of the enhancement mode MOSFET transistor provides an almost exact square law (see [9]). It is easily shown that squaring devices can be used to construct general polynomial operators of the type necessary to realize the Bernstein system. Thus it appears that the topic of polynomial system approximation may have a ready practicality in terms of microcircuit technology.

2. Weierstrass Approximation. As the technical work of this study deals primarily with the Weierstrass approximation it is useful to comment in somewhat more detail on the existing

literature. In this regard we cite first the original contribution of Weierstrass [10] whose fundamental result in contemporary form reads as follows.

Let f be an arbitrary continuous function on R , the real line. Let $D \subset R$ be an arbitrary compact set. Then for every $\epsilon > 0$ there exists a finite polynomial p , such that $\sup\{|f(x) - p(x)| : x \in D\} < \epsilon$. The Weierstrass result, over the years, has drawn the attention of several distinguished mathematicians including Frechet [11], Bernstein [12] and Stone [13] who investigated the relationship to power series expansions, constructive methods for finding the polynomial and extended the result to R^n among other things.

More recently Prenter [14] considered a real separable Hilbert space H , and showed that if $K \subset H$ is compact, $\epsilon > 0$ and f continuous on H then there exists a finite polynomial operator p , such that

$$\sup_{x \in K} ||f(x) - p(x)|| < \epsilon. \quad (2)$$

In a similar effort Prenter [15] and Ahmed [16] were able to use normed linear spaces.

The Prenter result [14], for example, states that if f is a continuous function on real $L_2(a,b)$ then on every compact subset there exists a finite number of kernels k_0, k_1, \dots, k_n such that p of equation 1 is an ϵ -approximation for f . Actually we would suspect more, namely, that if f is causal

then each kernel k_i is causal, that is for instance $k_1(t, \tau) = 0, \tau > t$. More generally, can a causality structure be superimposed on the function and its approximation? This question is answered affirmatively in [17] and [18].

The setting for [17] and [18] is a Hilbert resolution space¹ $\{H, P^t\}$ where H is real and separable. The set $K \subset H$ is always compact. The sets: C , SC , M , $C(K)$, and P denote the causal, strictly causal, memoryless, continuous on $K \subset H$, and polynomic functions, respectively, on $\{H, P^t\}$. For brevity we shall say that P is dense in $C(K)$ in the sense of Prenter's theorem.

The results of [17] include the following. The set $P \cap SC$ is dense in $C(K) \cap SC$. In L_2 the stronger result that $P \cap SC$ is dense in $C(K) \cap C$ is also established. This last result does not abstract. In ℓ_2 it is known [18] that $P \cap SC$ is not dense in $C(K) \cap C$.

All of the above results are nonconstructive in that

¹ A formal definition is given later.

they give no clue as to finding the polynomial approximate of a given function. On the real line, however, several constructive forms of the Weierstrass result do exist and the Bernstein polynomials constitute one of the more intriguing approaches to such constructions. In section (3) we develop a generalization of the Bernstein polynomials to real Hilbert space. Using a causal data interpolation scheme identification of the $p \in P \cap SC$ that approximates $f \in C(K) \cap C$ results.

In section (4) we consider the realization of the operator p in state variable form. This objective is also achieved as is evidenced by the following theorem of section (4).

Theorem. For every $f \in C(K) \cap C$ and $\epsilon > 0$, there exists a differential system

$$\begin{aligned}\dot{z}(t) &= A(t)z(t) + B(t)u(t) \quad , \quad z(0) = 0, \\ w(t) &= \phi(z(t), t)\end{aligned}$$

where $\phi((\cdot), t)$ is polynomial, such that the map $w = p(u)$ satisfies

$$\sup_K ||f(u) - p(u)|| < \epsilon.$$

For obvious reasons the equations of this theorem are called a Bernstein system.

3. The Bernstein Polynomials. We shall make use of a classical result on the Bernstein polynomials. It is helpful to introduce this result and associated notation before

proceeding. First we define the functions $p_n(v:(\cdot))$ on $[0,1]$ by the expression

$$p_n(v:x) = \binom{n}{v} x^v (1-x)^{n-v}, \quad x \in [0,1] \quad (3)$$

where $v \in \{0,1,\dots,n\}$.

Now let f be a bounded function on $[0,1]$. The polynomial $\hat{f}_n(x)$ is defined by the equation

$$\hat{f}_n(x) = \sum_{v=0}^n f(v/n) p_n(v:x), \quad x \in [0,1].$$

The familiar result of Bernstein is that $\hat{f}_n(x) \rightarrow f(x)$ at every point of continuity of f . Moreover, if f is continuous on $[0,1]$, then $\hat{f}_n \rightarrow f$ uniformly. While we shall not dwell on the proof of this result it is noted that positivity of $p_n(v:x)$ on $[0,1]$ was used by Bernstein in an essential way.

Consider then $\Omega = [0,1]^m \subset \mathbb{R}^m$ and let f be a function on \mathbb{R}^m which is bounded on Ω . Let $x = (x_1, \dots, x_m) \in \Omega$ and define the polymeric function

$$\hat{f}_n(x_1, \dots, x_m) = \sum_{v_1=0}^n \dots \sum_{v_m=0}^n f\left(\frac{v_1}{n}, \dots, \frac{v_m}{n}\right) p_n(v_1:x_1), \dots, p_n(v_m:x_m). \quad (4)$$

Using the single variate proof as a model it is easily shown that $\hat{f}_n \rightarrow f$ pointwise at all points of continuity of f and uniformly over Ω if f is continuous on Ω .

The farther extension that is needed in the present study is a relatively mild one. For this let H be any real Hilbert space and $\{e_1, \dots, e_m\}$ any finite set. We then form

$$\Omega = \{z = \sum_{i=1}^m x_i e_i : 0 \leq x_i \leq 1\} \quad (5)$$

and define the polynomic function \hat{f}_n by the equality (4). The proof for $H = \mathbb{R}^m$ and $\Omega = [0, 1]^m$ uses neither the linear independence nor the orthonormality of the coordinate basis of \mathbb{R}^m . As such the result; that $\|\hat{f}_n(x) - f(x)\| \rightarrow 0$ at all points of continuity of f and uniformly over Ω if f is continuous on Ω , follows by inspection.

To summarize the present adaptation of existing results to Hilbert spaces it is convenient to simplify notation. First let $\{e_1, \dots, e_m\}$ be linearly independent and let $\{e_1^+, \dots, e_m^+\}$ be the associated dual set; that is $\text{linear span}\{e_1, \dots, e_m\} = \text{linear span}\{e_1^+, \dots, e_m^+\}$ and $\langle e_i^+, e_j \rangle = \delta_{ij}$ $i, j = 1, \dots, m$. Let $\underline{n} = (n_1, n_2, \dots, n_m)$ and $\underline{v} = (v_1, v_2, \dots, v_m)$ be two integer tuplets. We shall use $\underline{n} \rightarrow \infty$ to denote the condition that all $n_i \rightarrow \infty$, although not necessarily at the same rate.

Using equation (3) we define

$$p(\underline{v}; x) = \prod_{i=1}^m p_{n_i}(v_i; \langle e_i^+, x \rangle). \quad (6)$$

We note that p is a scalar valued polynomic map defined on H . Using the notations

$$\sum_{\underline{v}=0}^{\underline{n}} = \sum_{v_1=0}^{n_1} \dots \sum_{v_m=0}^{n_m}$$

and

$$f(\underline{v}/\underline{n}) = f\left(\sum_{i=1}^m (v_i/n_i)e_i\right) \quad (7)$$

we define the polynomial function $\hat{f}_{\underline{n}}$ on H by the expression

$$\hat{f}_{\underline{n}}(x) = \sum_{\underline{v}=0}^{\underline{n}} f(\underline{v}/\underline{n}) p(\underline{v}:x), \quad x \in H. \quad (8)$$

The result which can be gleaned from the classical theory is the following

Theorem (1). If f is bounded on Ω then for each $x \in \Omega$, and $\epsilon > 0$ there is a finite tuple \underline{n} such that $||f(x) - \hat{f}_{\underline{n}}(x)|| < \epsilon$ at every continuity point of f . If f is continuous on Ω then $||\hat{f}_{\underline{n}}(x) - f(x)|| \rightarrow 0$ as $\underline{n} \rightarrow \infty$ uniformly over Ω .

Remark. In the above theorem Ω is given by equation (5). For the more general situation

$$\Omega = \left\{ \sum_{i=1}^m \alpha_i e_i : 0 \leq \alpha_i \leq M_i \right\},$$

a simple change of variables $f(\underline{v}/\underline{n}) = f\left(\sum_{i=1}^m (v_i/n_i) M_i e_i\right)$

preserves the validity of the theorem. If K is an arbitrary compact set then one can show that an affine map, τ , exists such that $\tau(K) \cap \Omega \neq \emptyset$. Moreover τ has an affine left inverse, τ^+ . Letting $g = f\tau^+$ we construct $\hat{g}_{\underline{n}}$ to fit g on Ω and then show that $\hat{f}_{\underline{n}} = \hat{g}_{\underline{n}}\tau$ fits f on K . These manipulations are straight forward but cluttersome and are omitted to save space.

Causality Structure. We now augment our real valued Hilbert space with a resolution of the identity. The resultant Hilbert resolution space has proved to be a natural setting for the study of system theoretic problems (see [17], [20], [21], [22]).

Let v denote an ordered set with minimal and maximal elements t_0, t_∞ respectively. The weakly closed set of orthoprojectors $\{P^t: t \in v\}$ is a resolution of the identity provided $P^{t_0} = 0$, $P^{t_\infty} = I$ and $P^t \geq P^\beta$ whenever $t \geq \beta$. Here $P^t \geq P^\beta$ denotes the condition $P^t P^\beta = P^\beta P^t = P^\beta$. We use also the notation $P_t = I - P^t$ $t \in v$ and $\Delta(t, \beta) = P^t - P^\beta = P^t P_\beta = P_\beta P^t$. Note that $\Delta(t, \beta)$ is an orthoprojector and that $\Delta(t, \beta) = 0$ all $\beta \geq t$.

Consider now a function f on H with $f(0)=0$. Following [17], [21], [22] we say that f is causal if and only if

$$P^t f = P^t f P^t, \quad \text{all } t \in v. \quad (9)$$

Now let $\{t_0, t_1, \dots, t_{N-1}, t_\infty\}$ be an ordered finite subset of v . For convenience set $P^j = P^{t_j}$ and $\Delta(j) = \Delta(j, j-1) = P_{j-1} P^j$. Using equation (9) it follows easily that $\Delta(j) f = \Delta(j) f P^j$ holds for causal f . Since $\sum_{j=1}^N \Delta(j) = I$ it follows that

$$f = \sum_{j=1}^N \Delta(j) f P^j$$

and taking the limit as the mesh $\{t_0, \dots, t_{N-1}, t_\infty\}$ is refined one arrives at a natural integral representation,

$$f = \int_v dP(s) f P^s$$

for f . This abbreviated discussion is expanded on in the references cited. We use it here for motivational purposes.

Our attention now returns to the earlier Bernstein problem with H being a Hilbert resolution space. The set $\{e_1, \dots, e_m\}$ is said to be well posed provided $\{P^t e_i : i=1, \dots, m\}$ is linearly

independent for every $t > t_0$. For convenience we assume that $\{e_1, \dots, e_m\}$ is well posed. We note that if

$$\Omega = \left\{ \sum_{i=1}^m \alpha_i e_i : 0 \leq \alpha_i \leq 1 \right\}$$

$$\Omega^t = \left\{ \sum_{i=1}^m \beta_i P^t e_i : 0 \leq \beta_i \leq 1 \right\}$$

then $\Omega^t = P^t(\Omega)$. It is also obvious that when $\{e_1, \dots, e_m\}$ is well posed then $\{e_1^+, \dots, e_m^+\}$ is well posed and the set $\{(P^t e_1)^+, \dots, (P^t e_m)^+\}$ is well defined for all $t > t_0$.

Using equation (3) we make the extended definition

$$p(\underline{v}; x; t) = \prod_{i=1}^m p_{n_i}(\underline{v}_i : \langle (P^t e_i)^+, x \rangle), \quad \text{all } x \in H, \quad (10)$$

where $t_0 \neq t_{ev}$. It can be easily shown that p is continuous* in t , and, of course, polynomial in x , moreover

$$P^t x = 0 \Rightarrow p(\underline{v}; x; \beta) = 0, \quad \text{all } \beta \leq t. \quad (11)$$

Recall now that $f(\underline{v}/\underline{n})$, of equation (7), for each $\underline{v}/\underline{n}$, is an element of H . Choosing now a partition for \underline{v} we form

* This property assumes that $\{P^t\}$ is continuous in the sense that $\lim_{\epsilon \rightarrow 0} \|P^t P_{t-\epsilon} x\| = 0$ all $t \in \mathcal{V}$.

$$\sum_{j=1}^N \Delta(j) f(\underline{v}/\underline{n}) p(\underline{v}; x; t_j). \quad (12)$$

We now take the limit over partition refinement

$$\int dP(s) f(\underline{v}/\underline{n}) p(\underline{v}; x; s) = \lim_{N \rightarrow \infty} \sum_{j=1}^N \Delta(j) f(\underline{v}/\underline{n}) p(\underline{v}; x; t_j)$$

where the limit converges by a result of Masani [23].

Since the set $\underline{v} \leq \underline{n}$ is finite we may sum the above function without disturbing the convergence. This results in the definition

$$f_{\underline{n}}^c(x) = \sum_{\underline{v}=0}^{\underline{n}} \int dP(s) f(\underline{v}/\underline{n}) p(\underline{v}; x; s). \quad (13)$$

We note first that $f_{\underline{n}}^c$ is causal, indeed

$$P^t f_{\underline{n}}^c(x) = \sum_{\underline{v}=0}^{\underline{n}} \int_{t_0}^t dP(s) f(\underline{v}/\underline{n}) p(\underline{v}; x; s)$$

however

$$p(\underline{v}; x; s) = p(\underline{v}; P^t \tilde{x}; s) \quad s \leq t$$

and hence $P^t f_{\underline{n}}^c = P^t f_{\underline{n}}^c P^t$, $t > t_0$. Suppose now that f is causal. Thus in equation (12)

$$\begin{aligned} \Delta(j) f(\underline{v}/\underline{n}) &= P_{j-1} P^j f(\underline{v}/\underline{n}) \\ &= \Delta(j) (f P^j)(\underline{v}/\underline{n}) \\ &= \Delta(j) f\left(\sum_{i=1}^m (v_i/n_i) P^j e_i\right), \end{aligned}$$

therefore

$$\Delta(j) f_{\underline{n}}^c(x) = \Delta(j) \left\{ \sum_{\underline{v}=0}^{\underline{n}} (f P^j)(\underline{v}/\underline{n}) p(\underline{v}; P^j x; t_j) \right\}$$

The bracketed quantity, however, is the Bernstein approximate for $\Delta(j)fP^j$ thus

$$\sum_{j=1}^N \Delta(j) f_{\underline{n}}^c \rightarrow \sum_{j=1}^N \Delta(j) f$$

with convergence in the spirit of theorem (1).

We summarize our development in the following theorem

Theorem (2). If causal f is bounded on Ω then for each $x \in \Omega$, and $\epsilon > 0$ there exists \underline{n} such that $||f(x) - f_{\underline{n}}^c(x)|| < \epsilon$ at every continuity point of f . If f is continuous on Ω then \underline{n} exists such that $\sup_{x \in \Omega} ||f(x) - f_{\underline{n}}^c(x)|| < \epsilon$.

4. Realization. The explicit nature of equation (13) and theorem (2) have answered the central theoretical questions in rather complete form. It is of interest, however, to give a realization of the causal Bernstein function, $f_{\underline{n}}^c$, in state variable form. For this we shall focus attention on the real $L_2[0, d]$, for $0 < d < \infty$ equipped with the usual inner product and take the resolution of the identity to the familiar truncation operators

$$(P^t x)(\beta) = \begin{cases} x(\beta) & \beta \leq t \\ 0 & \beta > t \end{cases}$$

In this last expression and elsewhere we slur over the distinction between functions and equivalence classes of functions as is the convention in dealing with the Lebesgue spaces.

Consider first the set $\{P^t e_i : i=1, \dots, m\}$ which is well posed. In earlier studies, [19], [20], the computation of the parameterized dual set $\{(P^t e_i)^+\}$ has been explored in some detail. The pertinent results are summarized in the following development. Following the earlier work define now the $m \times m$ symmetric matrices $N(t)$, $X(t)$, $\Pi(t)$, $\Lambda(t)$ by the equalities (here $t \in (0, \infty)$)

$$N_{ij}(t) = \langle P^t e_i, e_j \rangle ||P^t e_i||^{-1} ||P^t e_j||^{-1}$$

$$X_{ij}(t) = e_i^2(t) ||P^t e_i||^{-2} \delta_{ij}$$

$$\Pi_{ij}(t) = e_i(t) e_j(t) ||P^t e_i||^{-1} ||P^t e_j||^{-1}$$

$$\Lambda_{ij}(t) = ||P^t e_i||^{-1} \delta_{ij}$$

where $X(t)$, $\Lambda(t)$ are diagonal as indicated by the Kronecker term δ_{ij} .

Note that $N(t)$ is the Grammian matrix of a linearly independant set namely $\{P^t e_i / ||P^t e_i||\}$ and as such is invertable for all $t > 0$. It is shown in the references cited that $K(t) = N(t)^{-1}$ satisfies the Riccati equation

$$\dot{K}(t) = 1/2\{X(t)K(t) + K(t)X(t)\} - K(t)\Pi(t)K(t). \quad (14)$$

The equality $K(t') = N(t')^{-1}$ provides an initialization on this last equation for any $t' > 0$. In particular since $\lim_{t \rightarrow t_\infty} N(t) = \text{Grammian}\{e_i / ||e_i||\}$ it may prove convenient to solve equation (14) in reverse time.

For convenience we let

$$\begin{aligned}\underline{e}_t &= \text{column}(P^t e_1, \dots, P^t e_m) \\ \underline{e}_t^+ &= \text{column}((P^t e_1)^+, \dots, (P^t e_m)^+)\end{aligned}$$

A result of [20] is that

$$\underline{e}_t^+ = \Lambda(t)K(t)\Lambda(t)\underline{e}_t, \quad t > 0 \quad (15)$$

which is an explicit formula for \underline{e}_t^+ . From equations (10), (13) we see that $\langle (P^t e_i)^+, x \rangle$ is a key ingredient. Using (15) it can be easily seen that these functions are computed in vector form by the expression

$$\Psi(x, t) = \int_0^t \Lambda(t)K(t)\Lambda(t)\underline{e}(s)x(s)ds, \quad (16)$$

where

$$\begin{aligned}\underline{e}(s) &= \text{column}(e_1(s), \dots, e_m(s)) \\ \Psi(x, t) &= \text{column}(\langle (P^t e_1)^+, x \rangle, \dots, \langle (P^t e_m)^+, x \rangle).\end{aligned}$$

We are now in position to state the relevant result of [20].

Theorem (3). The equality $z(t) = \Psi(x, t)$ holds if and only if

$$\begin{aligned}\dot{p}(t) &= -1/2X(t)p(t) + \Lambda(t)\underline{e}(t)x(t), \quad t \in (0, \infty) \\ p(0) &= 0 \\ z(t) &= \Lambda(t)K(t)p(t), \quad t \in (0, \infty).\end{aligned}$$

It is noted that this theorem provides an explicit state variable realization of the requisite map, $x \rightarrow \Psi$.

In stating our final result it is helpful to utilize the change of variables $z(t) = \Lambda(t)K(t)p(t)$ and the identity $\dot{\Lambda} = (-1/2)X\Lambda$ which results in the corollary

Corollary. The equality $z(t) = \Psi(x, t)$ holds if and only if

$$\begin{aligned} \dot{z}(t) &= -\Lambda(t)K(t)\Lambda^{-1}(t)z(t) + \Lambda(t)K(t)\Lambda(t)\underline{e}(t)x(t), \quad t \geq 0 \\ z(0) &= 0. \end{aligned} \quad (17)$$

We turn now to the computation of the resolvent integral of equation (13). Consider first a fixed $y \in L_2(0, \infty)$ and a bounded continuous scalar valued function $m(\cdot)$. Using the notation $\Delta(j)$ of section (3) we have

$$\Delta(j)ym(j) = \begin{cases} m(t_j)y(\beta) & t_{j-1} \leq \beta \leq t_j \\ 0 & \text{otherwise} \end{cases}$$

thus

$$\lim_{|t_j - t_{j-1}| \rightarrow 0} \Delta(j)ym(j) = \begin{cases} m(t)y(t) & t = t_j \\ 0 & \text{otherwise.} \end{cases}$$

Clearly then on L_2 , $z = \int dP(s)ym(s)$ if and only if

$$z(t) = m(t)y(t) \quad t \in (0, \infty).$$

For convenience let $y_{\underline{v}} = f(\underline{v}/n)$. Recall also the scalar valued functions, $p(\underline{v}:r)$ on $[0, 1]^m$

$$p(\underline{v}:\underline{r}) = \prod_{i=1}^m p_{n_i}(v_i:r_i) \quad (r_1, \dots, r_m) \in [0,1]^m$$

We construct the map $\phi: \{0,1\}^m \rightarrow L_2$ by

$$\phi(\underline{r}:t) = \sum_{\underline{v}=0}^{\underline{n}} y_{\underline{v}}(t) p(\underline{v}:\underline{r}), \quad \underline{r} \in [0,1]^m$$

The above development has culminated in the theorem.

Theorem (4). $\omega = f_{\underline{n}}^C(x)$ if and only if

$$\dot{z}(t) = -\Lambda(t)K(t)\Pi(t)\Lambda^{-1}(t)z(t) + \Lambda(t)K(t)\Lambda(t)\underline{e}(t)x(t), \quad t > 0$$

$$z(0) = 0,$$

$$w(t) = \phi(z(t):t).$$

Example. In this example we let $H = L_2(0,1)$ and consider the power functions

$$x_j(t) = (1/j)t^j \quad j = 1, \dots, t \in [0,1].$$

It is readily verified that the set $\{P^t x_j: j = 1, \dots\}$ is linearly independent for any $t \in (0,1]$ and that

$$||P^t x_j||^2 = t^{2j+1}/j^2(2j+1), \quad j = 1, \dots \quad (18)$$

In particular $||x_j|| = 1/j\sqrt{2j+1}$.

Now let Γ be given by the defining equality

$$\Gamma = \{x = \sum_{j=1}^{\infty} \alpha_j x_j : |\alpha_j| \leq M\}$$

where $0 < M < \infty$. Using standard techniques it is not difficult to verify the following facts about Γ . First Γ is compact and second

$$\left\| \sum_{j=\ell+1}^{\infty} \alpha_j x_j \right\| < \sqrt{2M}/\sqrt{\ell}$$

For any $\varepsilon > 0$ we choose $\ell(\varepsilon) > 8M^2/\varepsilon^2$ in which case

$$\Omega_{\varepsilon} = \left\{ x: \sum_{j=1}^{\ell(\varepsilon)} \alpha_j x_j: |\alpha_j| \leq M \right\}$$

contains an $\varepsilon/2$ cover for Γ . Using continuity arguments we construct approximations on Ω_{ε} .

The set $\{P^t x_j: j=1, \dots, \ell\}$ is linearly independent for all $t > 0$ and hence the tools of theorem 4 apply. For convenience let us assume $0 \leq \alpha_j \leq 1$ and make the choice $e_j = x_j$, $j=1, \dots, \ell$.

We move then to theorem (3) and its corollary. Using the defining equality it is easily seen that N is constant, a property which was explored in detail in [19], and that all quantities of equation (17) can be explicitly determined. In fact,

$$N_{ij} = \sqrt{(2i+1)(2j+1)}/(i+j+1)$$

$$X_{ij} = (2j+1)\delta_{ij}/t$$

$$\Pi_{ij}(t) = \sqrt{(2i+1)(2j+1)}/t$$

$$\underline{e}(t) = \text{col}(t, \dots, t^{\ell}/\ell)$$

It is important to note that the state vector of equation 17 is of dimension ℓ . This dimension is fixed from this point forward and does not vary with n as the f_n^C approximations are constructed. Of course as ε is reduced $\ell(\varepsilon)$ increases.

TABLE 1

v/n	$f(v/\bar{n})(t)$	$P_{n_1}(v_1: z_1(t))$	$P_{n_2}(v_2: z_2(t))$	$P(v: z(t))$
(0,0)	0	$[1-z_1(t)]^2$	$[1-z_2(t)]^2$	$[1-z_1(t)]^2 [1-z_2(t)]^2$
(0,.5)	$1+.5t^2$	$[1-z_1(t)]^2$	$2z_2(t) [1-z_2(t)]$	$2[1-z_1(t)]^2 z_2(t) [1-z_2(t)]$
(0,1)	$4+t^2$	$[1-z_1(t)]^2$	$z_2^2(t)$	$[1-z_1(t)]^2 z_2^2(t)$
(.5,0)	.5t	$2z_1(t) [1-z_1(t)]$	$[1-z_2(t)]^2$	$2z_1(t) [1-z_1(t)] [1-z_2(t)]^2$
(.5,.5)	$1+.5t+.5t^2$	$2z_1(t) [1-z_1(t)]$	$2z_2(t) [1-z_2(t)]$	$4z_1(t) z_2(t) [1-z_1(t)] [1-z_2(t)]$
(.5,1)	$4+.5t+t^2$	$2z_1(t) [1-z_1(t)]$	$z_2^2(t)$	$2z_1(t) [1-z_1(t)] z_2^2(t)$
(1,0)	t	$z_1^2(t)$	$[1-z_2(t)]^2$	$z_1^2(t) [1-z_2(t)]^2$
(1,.5)	$1+t+.5t^2$	$z_1^2(t)$	$2z_2(t) [1-z_2(t)]$	$2z_1(t) z_2(t) [1-z_2(t)]$
(1,1)	$4+t+t^2$	$z_1^2(t)$	$z_2^2(t)$	$z_1^2(t) z_2^2(t)$

To carry forward in a very explicit manner let us assume $\ell=2$. Then

$$K = \begin{bmatrix} 16 & -4\sqrt{15} \\ -4\sqrt{15} & 16 \end{bmatrix}, \quad X = t^{-1} \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}, \quad \Pi = t^{-1} \begin{bmatrix} 3 & \sqrt{15} \\ \sqrt{15} & 5 \end{bmatrix}$$

Using equation 18 it is clear that $\Lambda(t)$ and hence the matrices of equation 17 are readily and explicitly determinable. Our attention turns to the construction of the polynomial state \rightarrow output map of theorem 4.

Our function set is of the form $x(t) = \alpha t + \beta t^2$ with $0 \leq \alpha, \beta \leq 1$ and we consider the function

$$[f(x)](t) = x(t) + |\ddot{x}(t)|^2$$

for $y = f(x)$ it follows that

$$y(t) = 4\beta^2 + \alpha t + \beta t^2 \quad (19)$$

Let us choose $\underline{n} = (2, 2)$ in which case \underline{v} can take on nine values $\leq \underline{n}$. Using the previous notation the following table is readily constructed

TABLE 1

v/n	$f(v/n)(t)$	$p_{n_1}(v_1: z_1(t))$	$p_{n_2}(v_2: z_2(t))$	$p(\underline{v}: z(t))$
(0,0)	0	$[1-z_2(t)]^2$	$[1-z_2(t)]^2$	$[1-z_1(t)]^2 [1-z_2(t)]^2$
(0,.5)	$1+.5t^2$	$[1-z_1(t)]^2$	$2z_2(t) [1-z_2(t)]$	$2[1-z_1(t)]^2 z_2(t) [1-z_2(t)]$
(0,1)	$4+t^2$	$[1-z_1(t)]^2$	$z_2^2(t)$	$[1-z_1(t)]^2 z_2^2(t)$
(.5,0)	$.5t$	$2z_1(t) [1-z_1(t)]$	$[1-z_2(t)]^2$	$2z_1(t) [1-z_1(t)] [1-z_2(t)]^2$
(.5,.5)	$1+.5t+.5t^2$	$2z_1(t) [1-z_1(t)]$	$2z_2(t) [1-z_2(t)]$	$4z_1(t) z_2(t) [1-z_1(t)] [1-z_2(t)]$
(.5,1)	$4+.5t+t^2$	$2z_1(t) [1-z_1(t)]$	$z_2^2(t)$	$2z_1(t) [1-z_1(t)] z_2^2(t)$
(1,0)	t	$z_1^2(t)$	$[1-z_2(t)]^2$	$z_1^2(t) [1-z_2(t)]^2$
(1,.5)	$1+t+.5t^2$	$z_1^2(t)$	$2z_2(t) [1-z_2(t)]$	$2z_1(t) z_2(t) [1-z_2(t)]$
(1,1)	$4+t+t^2$	$z_1^2(t)$	$z_2^2(t)$	$z_1^2(t) z_2^2(t)$

The function $\phi(z(t); t)$ is formed by taking the dot product of the $f(v/n)(t)$ and $p(\underline{v}: z(t))$ columns of table 1. The resultant expression does not add any additional clarity

so we will forego this here. To illustrate, however, we note that when $\alpha=0$, $0 \leq \beta \leq 1$ then $z_1(t)=0$ and $z_2(t)=\beta$. The above dot product produces

$$\begin{aligned}\phi(z(t), t) &= (1 + .5t^2)(2)(\beta)(1-\beta) + (4 + t^2)\beta^2 \\ &= 2\beta(1 + \beta) + \beta t^2\end{aligned}$$

The system, f , produces

$$f(x) = 4\beta^2 + \beta t^2$$

Similarly with $0 \leq \alpha \leq 1$ and $\beta=0$ it can be verified easily that

$$(fx)(t) = \phi(z(t), t) = \alpha t.$$

5. Remarks and Conclusions. The example of the last section demonstrates several properties which are not explicit in theorem 4. We have already noted that the size of the state space does not change with n . We see also that the output constraint map ϕ is a memoryless function of the state. Finally it is clear from table 1 that ϕ can be related to a bilinear form. In short we let $\bar{z} = (1, z_1, z_1^2, z_2, z_2^2)$ in which case there exist a matrix $E(t)$ such that

$$\phi(z(t), t) = \bar{z}^*(t)E(t)\bar{z}(t) \quad t > 0$$

The matrix E being taken from table 1 in the obvious way. This form is also valid in the general case.

It was noted in the introduction that our results have the characteristics of a sampling theorem for systems. The development does not emphasize this, however, in these results one can see f_n^c as a reconstruction of f from the samples $\{f(v/n) : v=0, 1, \dots, n\}$.

It should be noted that given a set of input-output pairs $\{(x_i, y_i)\}$ it is possible, by easier means (see [19]), to construct a causal polynomial map p , such that $p(x_i) = y_i$ $i=1, \dots, n$. The polynomial construction of [19] however will not necessarily uniformly approximate the continuous function generating the original data. In the same spirit the Bernstein system of section (4), while it approximates the function, does not reproduce exactly the original input-output measurements. (It comes ϵ close, of course.)

In section (2) it was pointed out that the results of [14, 15, 16, 17] use a separability assumption on the domain. The constructive methods of this study, however, do not seem to require this and it thus appears that the earlier results have been strengthened as well as put in constructive form.

As regards the analytic operator approach [5, 6, 7, 8] the following comparative remarks are appropriate. The system is assumed to be differentiable with derivatives used in modeling the $u \rightarrow x$ map. This procedure has the following attributes:

- (i) The expansion is valid on an input set $\|u\| \leq k$.
- (ii) Computing of derivatives of high order is necessary and f must be known to a comparable accuracy.
- (iii) A family of nonlinear state equations with a linear state \rightarrow output map realizes the power series expansion.
- (iv) No input-output information is utilized in the expansion.

In the same spirit the technique of sections 3,4 has the following characteristics:

- (i') The expansion is valid on a compact set.
- (ii') f need only be known to be continuous.
- (iii') A family of linear state equations with a nonlinear state-output map realizes the polynomial expansion.
- (iv') Only input-output information is utilized in the expansion.

The complexity of working with compact sets is offset by the computation of derivatives, thus attributes (i) and (ii) are computationally on a par with (i') and (ii'). Concerning (iii) and (iii') the merits of shifting the nonlinear behavior to the output constraint can be judged only in light of a specific application. Attributes (iv) and (iv') serve to crystalize the comparison. If the system is precisely known then the power series expansion methodology seems to have the advantage, although one might question expanding at all in this case. If the system is unknown then the methodology of this study can be utilized to mathematically model the system using external measurements.

Consider now the question of convergence rate. In the power series methodology convergence properties parallel those of the power series on the real line. Similarly the Bernstein system of theorem (4) is a linear operator composed with a memoryless Bernstein map and as such convergence properties of the Bernstein polynomials on R are inherited by the Bernstein system. Thus an accurate assessment for comparative convergence is obtained by looking at the situation on the real line.

We consider then the set $[0,1] \subset \mathbb{R}$, which satisfies both (i) and (i') as the domain of interest. For the function f on \mathbb{R} let $m(\cdot)$ denote the modulus of continuity of f defined by

$$m(\alpha) = \sup\{|f(x)-f(y)| : 0 \leq x, y \leq 1, |x-y| < \alpha\}.$$

For continuous f obviously $m(\alpha) \rightarrow 0$ as $\alpha \rightarrow 0$. We cite some standard results which go back to Jackson [25], Vallee-Poussin [26] and Popoviciu [27].

Here we use b_n to denote the Bernstein polynomial of order n .

(a) For each n

$$\sup_{[0,1]} |f(x) - b_n(x)| \leq (5/4)m(n^{-1/2})$$

If f is differentiable and m_1 is the modulus of continuity of f' then we have also

(b) For each n

$$\sup_{[0,1]} |f(x) - b_n(x)| \leq (3/4)n^{-1/2}m_1(n^{-1/2})$$

In closing I would like to acknowledge the careful review and constructive suggestions of () which have materially improved the paper.

REFERENCES

- [1] Volterra, V., Theory of Functionals, Dover Publications Inc., New York, 1959.
- [2] Bedrosian, E. and S.O. Rice, "The Output Properties of Volterra Systems Driven by Harmonic and Gaussian Inputs" Proc. IEEE, Vol. 59, pp. 1688-1707, 1971.
- [3] Parente, R.B., "Nonlinear Differential Equations and Analytic Systems Theory, SIAM J. Appl. Math. Vol. 18, pp. 41-66, 1970.
- [4] Van Trees, H.L., "Functional Techniques for the Analysis of the Nonlinear Behavior of Phase-Locked Loops," Proc. IEEE, Vol. 52, pp. 894-911.
- [5] Brockett, R.W., Volterra Series and Geometric Control Theory. Automatica, Vol. 12, 1976.
- [6] Bruni, C., DiPillo, G., Koch, G., "On the Mathematical Models of Bilinear Systems", Ricerche di Automatica, Vol. 2, pp. 11-26, 1976.
- [7] Krener, A.H. "Linearization and Bilinearization of Control Systems", Proc. 1974 Allerton Conf. on Cir. and Sys. Th., pp. 834-843, 1974.
- [8] Gilbert, E.G., "Volterra Series and the Response of Nonlinear Differential Systems: Trans. Conf. on Inf. Sci. and Systems, Johns Hopkins University, March 1976.
- [9] R.S. Cobbold, "Theory and Applications of Field-Effect Transistors," Wiley-Interscience, New York, 1970, Section 7.1.3.
- [10] K. Weierstrass, "Über die Analytische Darstellbarkeit Sogenannter Willkürlicher Funktionen Reeller Argumente," Math. Werke, III Bd., 1903.
- [11] M. Frechet, "Sur les Fonctions Continues," Ann. de l'Ecole Normale Sup., Third Series, Vol. 27, 1910.
- [12] S. Bernstein, Démonstration du Théorème de Weierstrass, fondée sur le calcul des probabilités, Com. Soc. Math., Kharkow, (2), 13, 1912-13.
- [13] M.H. Stone, "The Generalized Weierstrass Approximation Theorem," Math. Mag., Vol. 21, pp. 167-183, 1948.
- [14] P.M. Prenter, "A Weierstrass Theorem for Real Separable Hilbert Spaces," J. Approximation Theory, Vol. 3, No. 4 pp. 341-351, Dec. 1970.

- [15] P.M. Prenter, "A Weierstrass Theorem for Real Normed Linear Space," Bull. American Math. Soc., Vol. 75, pp. 860-862, 1969.
- [16] N.W. Ahmed, "An Approximation Theorem for Continuous Functions on L_p spaces," Univ. of Ottawa, Ottawa, Can., TR-73-18, Nov. 1973.
- [17] W.A. Porter, T.M. Clark, and R.M. DeSantis, "Causality Structure and the Weierstrass Theorem: J. Math. Anal. Appl., Vol. 52, No. 2, Nov. 1975.
- [18] W.A. Porter, "The Common Causality Structure of Multilinear Maps and their Multipower Forms," J. Math, Anal. Appl., Vol. 57, No. 3, March 1977.
- [19] W.A. Porter, "Data Interpolation, Causality Structure, and System Identification," Information and Control, Vol. 29, No. 3., Nov. 1975.
- [20] W.A. Porter, "Causal Realization from Input-Output pairs," SIAM J. Control, to appear 1976.
- [21] Saeks, R., "Resolution Space, Operators and Systems" Lecture Notes Springer-Verlag, New York, 1973.
- [22] Gohberg, I.Z., and Krein, M.G., "Theory of Volterra Operators in Hilbert space and Applications," Amer. Math. Soc., 24, 1970.
- [23] Masani, P., "Orthogonally Scattered Measures," Advan. Math. 2, pp. 61-117, Amer. Math. Soc. 1968.
- [24] Brockett, R.W., "Nonlinear Systems and Differential Geometry," Proceeding IEEE, Vol. 64, #1, Jan. 1976, pp. 61-72.
- [25] Jackson, D., "The Theory of Approximation," Amer. Math. Soc. Coll. Publ., Vol. 11, 1930, New York.
- [26] Vallée-Poussin, Ch. J. DeLa, "Lecons sur l'approximation," Paris 1919.
- [27] Popoviciu, T., "Sur l'approximation des fonctions convexes d'ordre supérieur," Mathematica (Cluj), 10, 1935, pp. 49-54.
- [28] Schetzen, M. "Measurement of the Kernals of a Nonlinear System of Finite Order, Int. J. Control, Vol. 1, March 1965, pp. 257-263.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

19 REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR - TR - 77 - 0369	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) APPROXIMATION BY BERNSTEIN SYSTEMS.	5. TYPE OF REPORT & PERIOD COVERED Interim Rept.	
7. AUTHOR(s) William A. Porter	6. PERFORMING ORG. REPORT NUMBER	
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Michigan Department of Electrical Engineering Ann Arbor, Michigan 48109	8. CONTRACT OR GRANT NUMBER(s) AFOSR 77-3160 new	
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, Washington, DC 20332	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A6	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) SEL-108	12. REPORT DATE Mar 1977	
	13. NUMBER OF PAGES 26	
	15. SECURITY CLASS. (of this report) UNCLASSIFIED	
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) 2304 A6		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) system theory, multilinearity, hilbert space, functional approximation		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Let K denote an arbitrary compact subset of L_2 . Let f be any causal continuous function on L_2 . Then there is a linear differential system: $\dot{x}(t) = A(t)x(t) + B(t)u(t)$, and a memoryless polynomial state to output map $y(t) = \phi(x(t))$ such that the system, f , thereby computed satisfies $\sup_{x \in K} \ f(x) - \hat{f}(x)\ < \epsilon$ where $\epsilon > 0$ is arbitrary. This and other results are developed.		

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)